

# Toward Designing Intelligent PDEs for Computer Vision: An Optimal Control Approach

Risheng Liu<sup>a</sup>, Zhouchen Lin<sup>b</sup>, Wei Zhang<sup>c</sup>, Kewei Tang<sup>a</sup>, Zhixun Su<sup>a</sup>

<sup>a</sup>*School of Mathematical Sciences, Dalian University of Technology, Dalian, China.*

<sup>b</sup>*Microsoft Research Asia, Beijing, China, e-mail: zhoulin@microsoft.com.*

<sup>c</sup>*Department of Information Engineering, The Chinese University of Hong Kong, China.*

---

## Abstract

Many computer vision and image processing problems can be posed as solving partial differential equations (PDEs). However, designing PDE system usually requires high mathematical skills and good insight into the problems. In this paper, we consider designing PDEs for various problems arising in computer vision and image processing in a lazy manner: *learning PDEs from real data via data-based optimal control*. We first propose a general intelligent PDE system which holds the basic translational and rotational invariance rule for most vision problems. By introducing a PDE-constrained optimal control framework, it is possible to use the training data resulting from multiple ways (ground truth, results from other methods, and manual results from humans) to learn PDEs for different computer vision tasks. The proposed optimal control based training framework aims at learning a PDE-based regressor to approximate the unknown (and usually nonlinear) mapping of different vision tasks. The experimental results show that the learnt PDEs can solve different vision problems reasonably well. In particular, we can obtain PDEs not only for problems that traditional PDEs work well but also for problems that PDE-based methods have never been tried before, due to the difficulty in describing those problems in a mathematical way.

*Keywords:* Optimal Control, PDEs, Computer Vision, Image Processing.

---

## **1. Introduction**

The wide applications of partial differential equations (PDEs) in computer vision and image processing can be attributable to two main factors [1]. First, PDEs in classical mathematical physics are powerful tools to describe, model, and simulate many dynamics such as heat flow, diffusion, and wave propagation. Second, many variational problems or their regularized counterparts can often be effectively solved from their Euler-Lagrange equations. Therefore, in general there are two types of methods for designing PDEs for vision tasks. For the first kind of methods, PDEs are written down directly (e.g., anisotropic diffusion [2], shock filter [3], based on some understandings on the properties of mathematical operators and the physical natures of the problems, and curve-evolution-based equations [4]). The second kind of methods basically define an energy functional and then derive the evolution equations by computing the Euler-Lagrange equation of the energy functional (e.g., total-variation-based variational methods [5][6][7]). In either way, people have to heavily rely on their intuition on the vision tasks. Therefore, traditional PDE-based methods require good mathematical skills when choosing appropriate PDE forms and predicting the final effect of composing related operators such that the obtained PDEs roughly meet the goals. If people do not have enough intuition on a vision task, they may have difficulty in acquiring effective PDEs. For example, although there has been much work on PDE-based image segmentation [8][9][10][11], the basic philosophy is always to follow the strong edges in the image and also require the edge contour to be smooth. Can we have a PDE system for objective detection (Fig. 1) that locates the object region



Figure 1: Can we design a PDE system which can detect the object of interest (e.g., the helicopter in the left image) and does not respond if the object is absent (e.g., the right image)?

if the object is present and does not respond if the object is absent? We believe that this is a big challenge to human intuition and is much more difficult than traditional segmentation tasks if a PDE-based method is required, because it is hard to describe an object class, which may have significant variation in shape, texture and pose. Without using additional information to judge the content, the existing PDEs for segmentation, e.g., [10], always output an “object region” for any non-constant image. In short, current PDE design methods greatly limit the applications of PDEs to a wider and more complex scope. This motivates us to explore whether we can acquire PDEs that are less artificial yet more powerful. In this paper, we give an affirmative answer to this question. We demonstrate that learning particular coefficients of a general intelligent PDE system from a given training data set might be a possible way of designing PDEs for computer vision in a lazy manner. Furthermore, borrowing this learning strategy from machine learning can generalize PDEs techniques for more complex vision problems.

Inspired by the electromagnetic field theory and Maxwell’s equations [12], we assume that the visual processing has two coupled evolutions in different scale spaces: one is in the *image scale space*, which controls the evolution of the output, and the other is in the *indicator scale space* that helps collect the global informa-

tion to guide the evolution in the image scale space. In this way, our general intelligent PDE system consists of two coupled evolutionary PDEs. Both PDEs are coupled equations between the image and indicator, up to their second order partial derivatives. Another key idea of our general intelligent PDE system is to assume that the PDEs that are sought could be written as combinations of “atoms” which satisfy the general properties of vision tasks. As a preliminary investigation, we utilize all the translational and rotational invariants as such “atoms” and propose the general intelligent PDE system as a linear combination of all these invariants [13]. Then the problem boils down to determining the combination coefficients among such “atoms”.

The theory of optimal control [14] has been well developed for over fifty years. With the enormous advances in computing power, optimal control is now widely used in multi-disciplinary applications such as biological systems, communication networks and socio-economic systems etc [15]. Optimal design and parameter estimation of systems governed by PDEs give rise to a class of problems known as PDE-constrained optimal control [16]. In this paper, a PDE-constrained optimal control technique as the training tool is introduced for our PDE system. We further propose a general framework for learning PDEs to accomplish a specific vision task via PDE-constrained optimal control, where the objective functional is to minimize the difference between the expected outputs and the actual outputs of the PDEs, given the input images. Such input-output image pairs are provided in multiple ways (e.g., ground truth, results from other methods or manually generated results by humans) for different tasks. Therefore, we can train the general intelligent PDE system to solve various vision problems which traditional PDEs may find difficult or even impossible.

In summary, our contributions are as follows:

1. Our intelligent PDE system provides a new way to design PDEs for computer vision. Based on this framework, we can design particular PDEs for different vision tasks using different sources of training images<sup>1</sup>. This may be very difficult for traditional PDE design methods. However, we would like to remind the readers that we have no intention to beat all the existing approaches for each task, because these approaches have been carefully and specially tuned for the task.
2. We propose a general data-based optimal control framework for training the PDE system. Fed with pairs of input and output images, the proposed PDE-constrained optimal control training model can automatically learn the combination coefficients in the PDE system. Unlike previous design methods, our approach requires much less human wits and can solve more difficult problems in computer vision.

The rest of the paper is structured as follows. We first introduce in Section 2 the general intelligent PDE system. In Section 3 we utilize the PDE-constrained optimal control technique as the training framework for our intelligent PDE system. Then in Section 4 we evaluate our intelligent PDE system with optimal control training framework by a series of computer vision and image processing problems. Finally, we give concluding remarks and a discussion on the future work in Section 5.

---

<sup>1</sup>Similar idea also appeared in [17]. But in that work, the authors only train special PDEs involving the curvature operator for basic image restoration tasks. In contrast, our work here proposes a more unified and elegant framework for more problems in computer vision.

## 2. General PDE System for Computer Vision

### 2.1. *Electromagnetic Field vs. Image Evolution*

Electromagnetism is the force that causes the interaction among electrically charged particles. The areas in which electromagnetic interaction happens are called the electromagnetic fields. But in physics, electrically charged objects were first thought to produce two types of fields associated with their charge property: An electric field and a magnetic field. Over time, it was realized that the electric and magnetic fields are better thought of as two parts of a greater whole – the electromagnetic field [12]. It affects the behavior of charged objects in the vicinity of the field. The electromagnetic field extends infinitely throughout space and describes the electromagnetic interaction. The theoretical implications of electromagnetism also led to the development of special relativity by Albert Einstein in 1905.

In this paper, inspired by this fundamental force of nature, we consider the image evolution in a similar way. For a target image signal  $u(t)$ , different from most traditional ways, which only consider the evolution in the image scale space, we define a companion signal  $v(t)$  named the indicator signal. It changes with time and guides the evolution of  $u(t)$  by collecting large scale information in the image. In this way, these two signals evolve in two coupled scale spaces.

### 2.2. *The Intelligent PDE System*

Similar to Maxwell’s equations [12], which are a set of PDEs describing how the electric and magnetic fields relate to their sources and how they develop with time, we propose a general PDE system for the evolution of our coupled signals.

The space of all PDEs is infinitely dimensional. To find the right form, we start with the properties that our PDE system should have, in order to narrow down the search space. We notice that translationally and rotationally invariant properties are very important for computer vision, i.e., in most vision tasks, when the input image is translated or rotated, the output image is also translated or rotated by the same amount. So we require that our PDE system is translationally and rotationally invariant. According to the differential invariant theory [13], the form of our PDEs must be functions of the fundamental differential invariants under the group of translation and rotation. The fundamental differential invariants are invariant under translation and rotation and other invariants can be written as their functions. We list those up to second order in Table 1, where some notations can be found in Table 2. In the sequel, we shall use  $\{\text{inv}_j(u, v)\}_{j=0}^{16}$  to refer to them in order. Note that those invariants are ordered with  $u$  going before  $v$ . We may reorder them with  $v$  going before  $u$ . In this case, the  $j$ -th invariant will be referred to as  $\text{inv}_j(v, u)$ . So the simplest choice of our general PDE system is the linear combination of the differential invariants, leading to the following form:

$$\left\{ \begin{array}{ll} \frac{\partial u}{\partial t} - F(u, v, \{a_j\}_{j=0}^{16}) = 0, & (x, y, t) \in Q, \\ u(x, y, t) = 0, & (x, y, t) \in \Gamma, \\ u|_{t=0} = f_u, & (x, y) \in \Omega, \\ \frac{\partial v}{\partial t} - F(v, u, \{b_j\}_{j=0}^{16}) = 0, & (x, y, t) \in Q, \\ v(x, y, t) = 0, & (x, y, t) \in \Gamma, \\ v|_{t=0} = f_v, & (x, y) \in \Omega, \end{array} \right. \quad (1)$$

where

$$\begin{aligned} F(u, v, \{a_j\}_{j=0}^{16}) &= \sum_{j=0}^{16} a_j(t) \text{inv}_j(u, v), \\ F(v, u, \{b_j\}_{j=0}^{16}) &= \sum_{j=0}^{16} b_j(t) \text{inv}_j(v, u), \end{aligned} \quad (2)$$

$\Omega$  is the rectangular region occupied by the input image  $I$ ,  $T$  is the time that the PDE system finishes the visual information processing and outputs the results, and  $f_u$  and  $f_v$  are the initial functions of  $u$  and  $v$ , respectively. The meaning of other notations in (1) can be found in Table 2. For computational issues and the ease of mathematical deduction,  $I$  will be padded with zeros of several pixels width around it. As we can change the unit of time, it is harmless to fix  $T = 1$ .  $\{a_j(t)\}_{j=0}^{16}$  and  $\{b_j(t)\}_{j=0}^{16}$  are sets of functions defined on  $Q$  that are used to control the evolution of  $u$  and  $v$ , respectively. As  $\nabla u$  and  $\mathbf{H}_u$  change to  $\mathbf{R}\nabla u$  and  $\mathbf{R}\mathbf{H}_u\mathbf{R}^T$ , respectively, when the image is rotated by a matrix  $\mathbf{R}$ , it is easy to check the rotational invariance of those quantities. So the PDE system (1) is rotationally invariant. Furthermore, the following proposition implies that the control functions  $a_j(t)$  and  $b_j(t)$  can be functions of  $t$  only.

**Proposition 2.1.** *Suppose the PDE system (1) is translationally invariant, then the control functions  $\{a_j\}_{j=0}^{16}$  and  $\{b_j\}_{j=0}^{16}$  must be independent of  $(x, y)$ .*

The proof of Proposition 2.1 is presented in Appendix A.

### 3. Training the PDE System via Data-based Optimal Control

In this section, we propose a data-based optimal control framework to train the intelligent PDE system for particular vision tasks.

#### 3.1. The Objective Functional

Given the forms of PDEs shown in (1), we have to determine the coefficient functions  $a_j(t)$  and  $b_j(t)$ . We may prepare training samples  $\{(I_m, O_m)\}_{m=1}^M$ , where  $I_m$  is the input image and  $O_m$  is the expected output image, and compute



Table 1: Translationally and rotationally invariant fundamental differential invariants up to the second order.

$j$	$\text{inv}_j(u, v)$
0,1,2	$1, v, u$
3,4	$\ \nabla v\ ^2 = v_x^2 + v_y^2, \ \nabla u\ ^2 = u_x^2 + u_y^2$
5	$(\nabla v)^T \nabla u = v_x u_x + v_y u_y$
6,7	$\text{tr}(\mathbf{H}_v) = v_{xx} + v_{yy}, \text{tr}(\mathbf{H}_u) = u_{xx} + u_{yy}$
8	$(\nabla v)^T \mathbf{H}_v \nabla v = v_x^2 v_{xx} + 2v_x v_y v_{xy} + v_y^2 v_{yy}$
9	$(\nabla v)^T \mathbf{H}_u \nabla v = v_x^2 u_{xx} + 2v_x v_y u_{xy} + v_y^2 u_{yy}$
10	$(\nabla v)^T \mathbf{H}_v \nabla u = v_x u_x v_{xx} + (v_x u_y + v_y u_x) v_{xy} + v_y u_y v_{yy}$
11	$(\nabla v)^T \mathbf{H}_u \nabla u = v_x u_x u_{xx} + (v_x u_y + v_y u_x) u_{xy} + v_y u_y u_{yy}$
12	$(\nabla u)^T \mathbf{H}_v \nabla u = u_x^2 v_{xx} + 2u_x u_y v_{xy} + u_y^2 v_{yy}$
13	$(\nabla u)^T \mathbf{H}_u \nabla u = u_x^2 u_{xx} + 2u_x u_y u_{xy} + u_y^2 u_{yy}$
14	$\text{tr}(\mathbf{H}_v^2) = v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2$
15	$\text{tr}(\mathbf{H}_v \mathbf{H}_u) = v_{xx} u_{xx} + 2v_{xy} u_{xy} + v_{yy} u_{yy}$
16	$\text{tr}(\mathbf{H}_u^2) = u_x^2 + 2u_{xy}^2 + u_y^2$

Table 2: Notations.

$\Omega$	An open bounded region in $\mathbb{R}^2$	$\partial\Omega$	Boundary of $\Omega$
$(x, y)$	$(x, y) \in \Omega$ , spatial variable	$t$	$t \in (0, T)$ , temporal variable
$Q$	$\Omega \times (0, T)$	$\Gamma$	$\partial\Omega \times (0, T)$
$ \cdot $	The area of a region	$\mathbf{X}^T$	Transpose of matrix (or vector)
$\ \cdot\ $	$L^2$ norm	$\text{tr}(\cdot)$	Trace of matrix
$\nabla u$	Gradient of $u$	$\mathbf{H}_u$	Hessian of $u$
$\wp$	$\wp = \{(0, 0), (0, 1), (1, 0), (0, 2), (1, 1), (2, 0)\}$ , index set for partial differentiation		

the coefficient functions that minimize the following functional:

$$\begin{aligned}
& J(\{u_m\}_{m=1}^M, \{a_j\}_{j=0}^{16}, \{b_j\}_{j=0}^{16}) \\
&= \frac{1}{2} \sum_{m=1}^M \int_{\Omega} [u_m(x, y, 1) - O_m]^2 d\Omega \\
&+ \frac{1}{2} \sum_{j=0}^{16} \lambda_j \int_0^1 a_j^2(t) dt + \frac{1}{2} \sum_{j=0}^{16} \mu_j \int_0^1 b_j^2(t) dt,
\end{aligned} \tag{3}$$

where  $u_m(x, y, 1)$  is the output image at time  $t = 1$  computed from (1) when the input image is  $I_m$ , and  $\lambda_i$  and  $\mu_i$  are positive weighting parameters. The first term requires that the final output of our PDE system be close to the ground truth. The second and the third terms are for regularization so that the optimal control problem is well posed, as there may be multiple minimizers for the first term.

### 3.2. Solving the Optimal Control Problem

Then we have the following optimal control problem with PDE constraints:

$$\begin{aligned} & \arg \min_{\{a_j\}_{j=0}^{16}, \{b_j\}_{j=0}^{16}} J(\{u_m\}_{m=1}^M, \{a_j\}_{j=0}^{16}, \{b_j\}_{j=0}^{16}) \\ & \text{s.t.} \begin{cases} \frac{\partial u}{\partial t} - F(u, v, \{a_j\}_{j=0}^{16}) = 0, & (x, y, t) \in Q, \\ u(x, y, t) = 0, & (x, y, t) \in \Gamma, \\ u|_{t=0} = f_u, & (x, y) \in \Omega, \\ \frac{\partial v}{\partial t} - F(v, u, \{b_j\}_{j=0}^{16}) = 0, & (x, y, t) \in Q, \\ v(x, y, t) = 0, & (x, y, t) \in \Gamma, \\ v|_{t=0} = f_v, & (x, y) \in \Omega. \end{cases} \end{aligned} \quad (4)$$

By introducing the adjoint equation of (4), the Gâteaux derivative of  $J$  can be computed and consequently, the (local) optimal  $\{a_j\}_{j=0}^{16}$  and  $\{b_j\}_{j=0}^{16}$  can be computed via gradient-based algorithms (e.g., conjugate gradient). Here, we give the adjoint equation and Gâteaux derivative directly:

#### 3.2.1. Adjoint Equation

$$\begin{aligned} & \begin{cases} \frac{\partial \varphi_m}{\partial t} + E(u_m, v_m, \varphi_m, \phi_m) = 0, & (x, y, t) \in Q, \\ \varphi_m = 0, & (x, y, t) \in \Gamma, \\ \varphi_m|_{t=1} = O_m - u_m(1), & (x, y) \in \Omega, \\ \frac{\partial \phi_m}{\partial t} + E(v_m, u_m, \phi_m, \varphi_m) = 0, & (x, y, t) \in Q, \\ \phi_m = 0, & (x, y, t) \in \Gamma, \\ \phi_m|_{t=1} = 0, & (x, y) \in \Omega, \end{cases} \end{aligned} \quad (5)$$

where

$$\begin{aligned}
& E(u_m, v_m, \varphi_m, \phi_m) \\
&= \sum_{(p,q) \in \wp} (-1)^{p+q} \frac{\partial^{p+q} (\sigma_{pq}(u_m) \varphi_m + \sigma_{pq}(v_m) \phi_m)}{\partial x^p \partial y^q}, \\
& E(v_m, u_m, \phi_m, \varphi_m) \\
&= \sum_{(p,q) \in \wp} (-1)^{p+q} \frac{\partial^{p+q} (\sigma_{pq}(u_m) \varphi_m + \sigma_{pq}(v_m) \phi_m)}{\partial x^p \partial y^q}, \\
& \sigma_{pq}(u) = \frac{\partial F(u)}{\partial u_{pq}} = \sum_{j=0}^{16} a_j \frac{\partial \text{inv}_j(u, v)}{\partial u_{pq}}, \quad u_{pq} = \frac{\partial^{p+q} u}{\partial x^p \partial y^q}, \\
& \sigma_{pq}(v) = \frac{\partial F(v)}{\partial v_{pq}} = \sum_{j=0}^{16} b_j \frac{\partial \text{inv}_j(v, u)}{\partial v_{pq}}, \quad v_{pq} = \frac{\partial^{p+q} v}{\partial x^p \partial y^q}.
\end{aligned} \tag{6}$$

### 3.2.2. Gâteaux Derivative of the Functional

With the help of the adjoint equation, at each iteration the derivative of  $J$  with respect to  $a_j(t)$  and  $b_j(t)$  are as follows:

$$\begin{aligned}
\frac{\partial J}{\partial a_j} &= \lambda_j a_j - \sum_{m=1}^M \int_{\Omega} \varphi_m \text{inv}_j(u_m, v_m) d\Omega, \quad j = 0, \dots, 16, \\
\frac{\partial J}{\partial b_j} &= \mu_j b_j - \sum_{m=1}^M \int_{\Omega} \phi_m \text{inv}_j(v_m, u_m) d\Omega, \quad j = 0, \dots, 16,
\end{aligned} \tag{7}$$

where the adjoint functions  $\varphi_m$  and  $\phi_m$  are the solutions to (5).

### 3.2.3. Initialization

Good initialization increases the approximation accuracy of the learnt PDEs. In our current implementation, we simply set the initial functions of  $u$  and  $v$  as the input image:

$$u_m(x, y, 0) = v_m(x, y, 0) = I_m(x, y), \quad m = 1, 2, \dots, M.$$

Then we employ a heuristic method to initialize the control functions. At each time step,  $\frac{\partial u_m}{\partial t}$  is expected to be  $d_m(t) = \frac{O_m - u_m(t)}{1-t}$  so that  $u_m(t)$  moves towards the expected output  $O_m$  and by the form of (1) we may solve  $\{a_j(t)\}_{j=0}^{16}$  such that

$$\sum_{m=1}^M \int_{\Omega} [F(u_m, v_m, \{a_j(t)\}_{j=0}^{16}) - d_m(t)]^2 d\Omega \tag{8}$$

is minimized<sup>2</sup>. In this way, we initialize  $a_j(t)$  successively in time while fixing  $b_j(t) = 0, j = 0, 1, \dots, 16$ .

#### 3.2.4. Finite Difference Method for Numerical Solution

To solve the intelligent PDE system numerically, we design a finite difference scheme [18] for the PDEs. We discretize the PDEs, i.e. replace the derivatives  $\frac{\partial f}{\partial t}$ ,  $\frac{\partial f}{\partial x}$  and  $\frac{\partial^2 f}{\partial x^2}$  with finite differences as follows:

$$\begin{cases} \frac{\partial f}{\partial t} = \frac{f(t+\Delta t) - f(t)}{\Delta t}, \\ \frac{\partial f}{\partial x} = \frac{f(x+1) - f(x)}{2}, \\ \frac{\partial^2 f}{\partial x^2} = f(x-1) - 2f(x) + f(x+1). \end{cases} \quad (9)$$

The discrete forms of  $\frac{\partial f}{\partial y}$ ,  $\frac{\partial^2 f}{\partial y^2}$  and  $\frac{\partial^2 f}{\partial x \partial y}$  can be defined similarly. In addition, we discretize the integrations as

$$\begin{cases} \int_{\Omega} f(x, y) d\Omega = \frac{1}{N} \sum_{\Omega} f(x, y), \\ \int_0^t f(t) dt = \Delta t \sum_{i=0}^{T_m} f(i \cdot \Delta t), \end{cases} \quad (10)$$

where  $N$  is the number of pixels in the spatial area,  $\Delta t$  is a properly chosen time step size and  $T_m = \lfloor \frac{1}{\Delta t} + 0.5 \rfloor$  is the index of the expected output time. Then we use an explicit scheme to compute the numerical solutions.

#### 3.3. The Optimal-Control-Based Training Framework

We now summarize in Algorithm 1 the data-based optimal control training framework for the intelligent PDE system. After the PDE system is learnt, it can be applied to new test images by solving (1), whose inputs  $f_u$  and  $f_v$  are both the test image and the solution  $u(t)|_{t=1}$  is the desired output image.

---

<sup>2</sup>It is to minimize the difference between the left and the right hand sides of (1).

---

**Algorithm 1 (Data-based optimal control framework for training the PDE system)**

---

**Input:** Training image pairs  $\{(I_m, O_m)\}_{m=1}^M$ .

- 1: Initialize  $a_j(t)$ ,  $t = 0, \Delta t, \dots, 1 - \Delta t$ , by minimizing (8) and fix  $b_j(t) = 0$ ,  $j = 0, 1, \dots, 16$ .
- 2: **while** not converged **do**
- 3:   Compute  $\frac{\partial J}{\partial a_j}$  and  $\frac{\partial J}{\partial b_j}$ ,  $j = 0, \dots, 16$ , using (7).
- 4:   Decide the search direction using the conjugate gradient method [19].
- 5:   Perform golden search along the search direction and update  $a_j(t)$  and  $b_j(t)$ ,  $j = 0, \dots, 16$ .
- 6: **end while**

**Output:** The coefficient functions  $\{a_j(t)\}_{j=0}^{16}$  and  $\{b_j(t)\}_{j=0}^{16}$ .

---

## 4. Experimental Results

In this section, we apply our data-based optimal control framework to learn PDEs for four groups of basic computer vision problems: Natural image denoising, edge detection, blurring and deburring, and image segmentation and object detection. As our goal is to show that the data-based optimal control framework could be a new approach for designing PDEs and an effective regressor for many computer vision tasks, *NOT* to propose better algorithms for these tasks, we are not going to fine tune our PDEs and then compare it with the state-of-the-art algorithms in every task.

### 4.1. Learning from Ground Truth: Natural Image Denoising

Image denoising is one of the most fundamental low-level vision problems. For this task, we compare our learnt PDEs with the existing PDE-based denoising

methods, ROF [5] and TV- $L^1$  [6], on images with unknown natural noise. This task is designed to demonstrate that our method can solve problems by learning from the ground truth. This is the first advantage of our data-based optimal control model. We take 240 images, each with a size of  $150 \times 150$  pixels, of 11 objects using a Canon 30D digital camera, setting its ISO to 1600. For each object, 30 images are taken without changing the camera settings (by fixing the focus, aperture and exposure time) and without moving the camera position. The average image of them can be regraded as the noiseless ground truth image. We randomly choose 8 objects. For each object we randomly choose 5 noisy images. These noisy images and their ground truth images are used to train the PDE system. Then we compare our learnt PDEs with the traditional PDEs in [5] and TV- $L^1$  [6] on images of the remaining 3 objects.

Fig. 2 shows the comparison results. One can see that the PSNRs of our intelligent PDEs are dramatically higher than those of traditional PDEs. This is because our data-based PDE learning framework can easily adapt to unknown types of noise and obtain PDE forms to fit for the natural noise well, while most traditional PDE-based denoising methods were designed under specific assumptions on the types of noise (e.g., ROF is designed for Gaussian noise [5] while TV- $L^1$  is designed for impulsive noise [20]). Therefore, they may not fit for unknown types of noise as well as our intelligent PDEs. The curves of the learnt coefficients for image denoising are shown in Fig. 3.

#### 4.2. *Learning from Other Methods: Edge Detection*

The image edge detection task is used to demonstrate that our PDEs can be learnt from the results of different methods and achieve a better performance than all of them. This is another advantage of our data-based optimal control model.

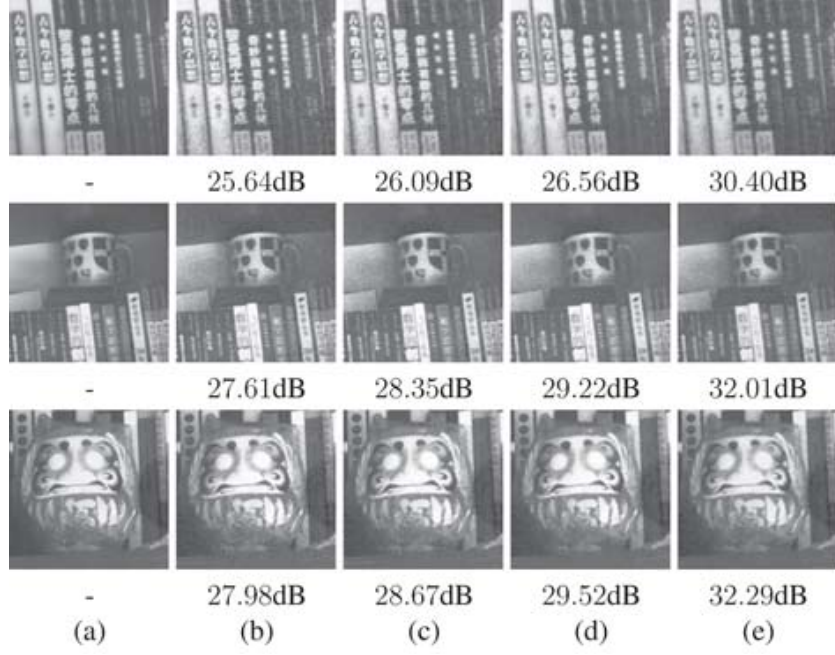


Figure 2: The results of denoising images with natural noise. (a) original noiseless image. (b) noisy image with real noise. (c)-(e) denoised images using the ROF,  $TV-L^1$ , and our intelligent PDE system, respectively. The PSNRs are presented below each image.

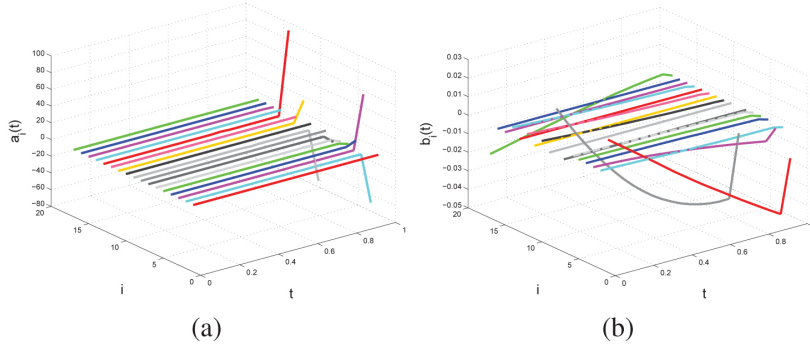


Figure 3: Learnt coefficients of the intelligent PDE system for the natural noise denoising problem. (a) the coefficients  $\{a_j(t)\}_{j=0}^{16}$  for the image evolution. (b) the coefficients  $\{b_j(t)\}_{j=0}^{16}$  for the indicator evolution.



For this task, we use three simple first order edge detectors [21] (Sobel, Roberts Cross, and Prewitt) to generate the training data. We randomly choose 7 images from the Berkeley image database [22] and use the above three detectors to generate the output images<sup>3</sup>, together with the input images, to train our PDE system for edge detection.

Fig. 4 shows part of the edge detection results on other images in the Berkeley image database. One can see that our PDEs respond selectively to edges and basically produce visually significant edges, while the edge maps of other three detectors are more chaotic. Note that the solution to our PDEs is supposed to be a more or less smooth function. So one cannot expect that our PDEs produce an exactly binary edge map. Instead, an approximation of a binary edge map is produced. The curves of the learnt coefficients for edge detection are shown in Fig. 5.

#### 4.3. *Learning to Solve Both Primal and Inverse Problems: Blurring and Deblurring*

The traditional PDEs for solving different problems are usually of very different appearance. The task of solving both blurring and deblurring is designed to show that the same form of PDEs can be learnt to solve both the primal and inverse problems. This is the third advantage of our data-based optimal control model.

For the image blurring task (the primal problem), the output image is the convolution of the input image with a Gaussian kernel. So we generate the output

---

<sup>3</sup>This implies that we actually use a kind of combination of the results from different methods to train our PDE system.

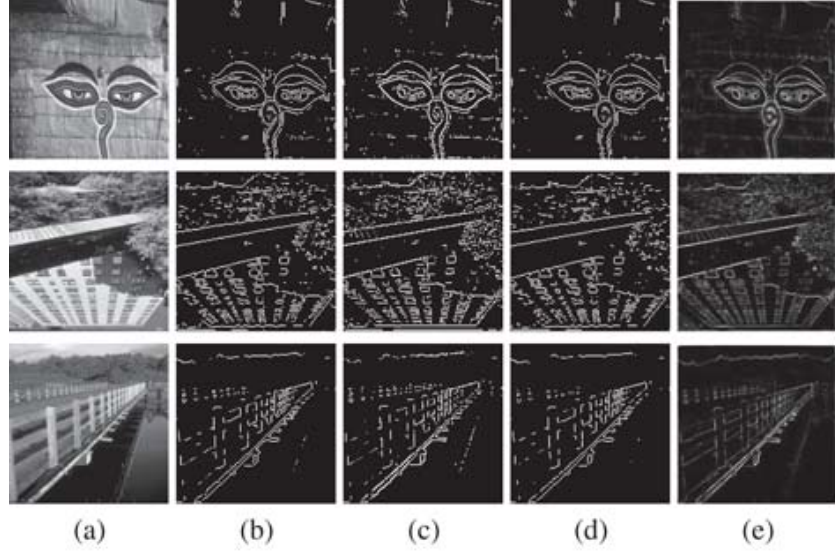


Figure 4: The results of edge detection. (a) original image. (b)-(e) edge detection results using the Sobel, Roberts Cross, Prewitt, and our intelligent PDE, respectively.

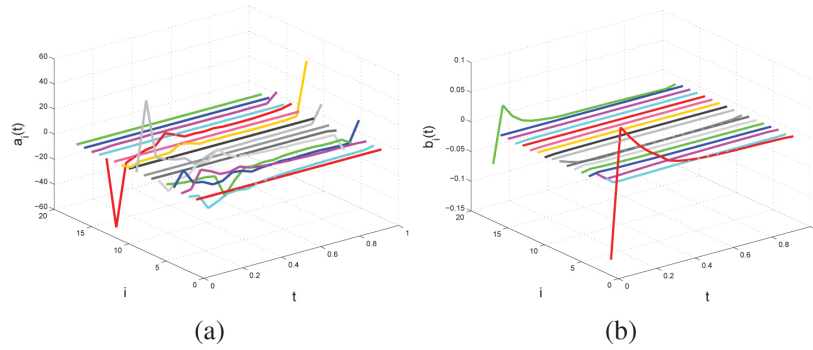


Figure 5: Learnt coefficients of the intelligent PDE system for the edge detection problem. (a) the coefficients  $\{a_j(t)\}_{j=0}^{16}$  for the image evolution. (b) the coefficients  $\{b_j(t)\}_{j=0}^{16}$  for the indicator evolution.

images by blurring high resolution images using a Gaussian kernel with  $\sigma = 1$ . The original images are used as the input. As shown in the third row of Fig. 6, the output is nearly identical to the ground truth (the second row of Fig. 6). For the image deblurring task (the inverse problem), we just exchange the input and output images for training. One can see in the bottom row of Fig. 6 that the output is very close to the original image (first row of Fig. 6). The curves of the learnt coefficients for image deblurring are shown in Fig. 7.

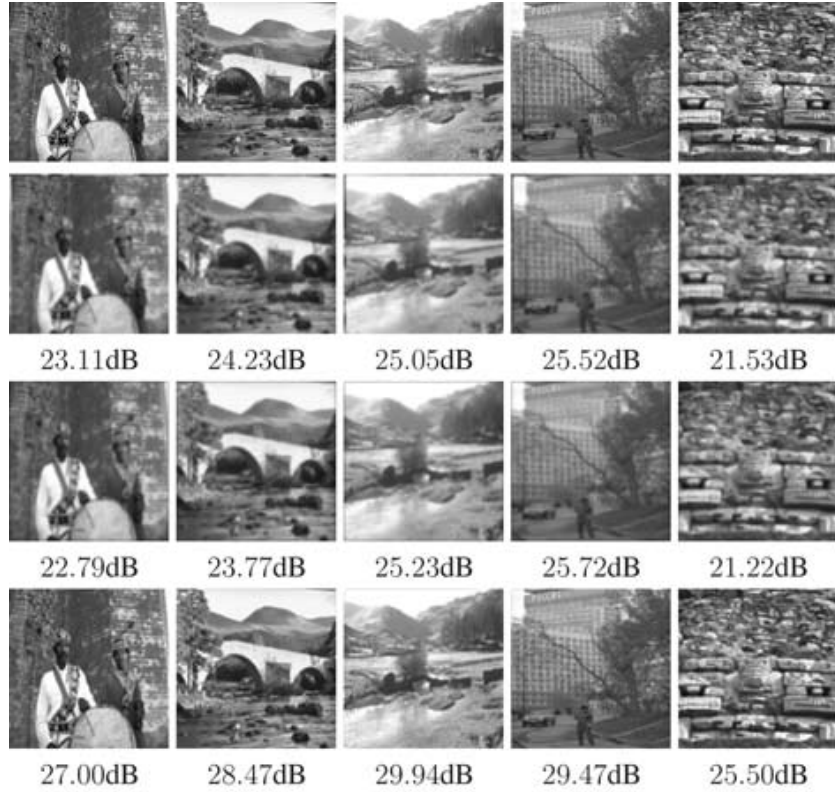


Figure 6: The results of image blurring and deblurring. The top row are the original images. The second row are the blurring results of a Gaussian kernel. The third row are the blurring results of our Intelligent PDE. The bottom row are the deblurring results of our Intelligent PDE. The PSNRs are presented below each image.

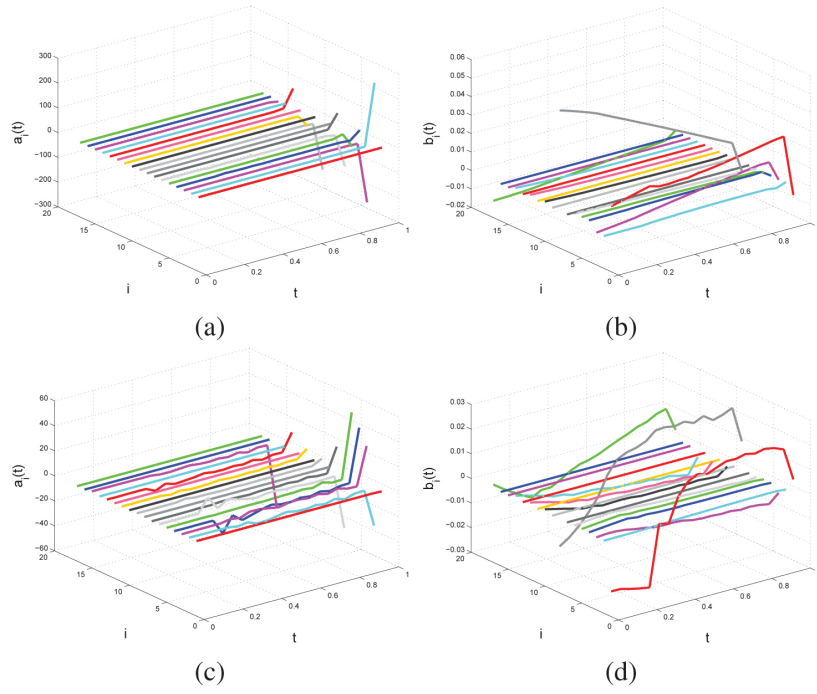


Figure 7: Learnt coefficients of the intelligent PDE system for image blurring (top row) and deblurring (bottom row) problems. (a) and (c) are the coefficients  $\{a_j(t)\}_{j=0}^{16}$  for the image evolutions. (b) and (d) are the coefficients  $\{b_j(t)\}_{j=0}^{16}$  for the indicator evolutions.

#### 4.4. *Learning from Humans: Image Segmentation and Object Detection*

Image segmentation and object detection are designed to demonstrate that our PDE system can learn from the human behavior directly (learn the segmentation and detection results provided by humans, e.g., manually segmented masks).

For image segmentation, it is a highly ill-posed problem and there are many criteria that define the goal of segmentation, e.g., breaking an image into regions with similar intensity, color, texture, or expected shape. As none of the current image segmentation algorithms can perform object level segmentation well out of complex backgrounds, we choose to require our PDEs to achieve a reasonable goal, namely segmenting relatively darker objects against relatively simple backgrounds, where both the foreground and the background can be highly textured and simple thresholding cannot separate them. So we select 60 images from the Corel image database [23] that have relatively darker foregrounds and relatively simple backgrounds, but the foreground is not of uniformly lower graylevels than the background, and also prepare the manually segmented binary masks as the outputs of the training images, where the black regions are the backgrounds (Fig. 8).



Figure 8: Examples of the training images for image segmentation. In each group of images, on the left is the input image and on the right is the ground truth output mask.

Part of the segmentation results are shown in Fig. 9, where we have set a threshold for the output mask maps of our learnt PDEs with a constant 0.5. We see that our learnt PDEs produce fairly good object masks. We also test the active

contour method by Li et al. [10]<sup>4</sup> and the normalized cut method [24]<sup>5</sup>. One can see from Fig. 9 that the active contour method cannot segment object details due to the smoothness constraint on the object shape and the normalized cut method cannot produce a closed foreground region. To provide a quantitative evaluation, we use the  $F$ -measures that merge the precision and recall of segmentation:

$$F_\alpha = \frac{(1 + \alpha) \cdot recall \cdot precision}{\alpha \cdot precision + recall}, \text{ where}$$

$$recall = \frac{|A \cap B|}{|A|}, \quad precision = \frac{|A \cap B|}{|B|},$$

in which  $A$  is the ground truth mask and  $B$  is the computed mask. The most common choice of  $\alpha$  is 2. On our test images, the  $F_2$  measures of our PDEs, [10] and [24] are  $0.90 \pm 0.05$ ,  $0.83 \pm 0.16$  and  $0.61 \pm 0.20$ , respectively. One can see that the performance of our PDEs is better than theirs, in both visual quality and quantitative measure. The curves of the learnt coefficients for image segmentation are shown in Fig. 10.

We also present the evolution process of the mask maps across time (Fig. 11). One can see that although the foreground is relatively darker than the background, the PDEs correctly detect the most salient points/edges and then propagate the information across the foreground region, resulting in a *brighter* output region for the foreground.

Now we apply our intelligent PDEs for a more complex task: Object detection. Namely, the PDEs should respond strongly to the object of interest while not responding (or responding much more weakly) if the object is absent in the image. It should be challenging enough for one to manually design PDEs to perform

---

<sup>4</sup>Code available at <http://www.engr.uconn.edu/~cmli/>

<sup>5</sup>Code available at <http://www.cis.upenn.edu/~jshi/software/>

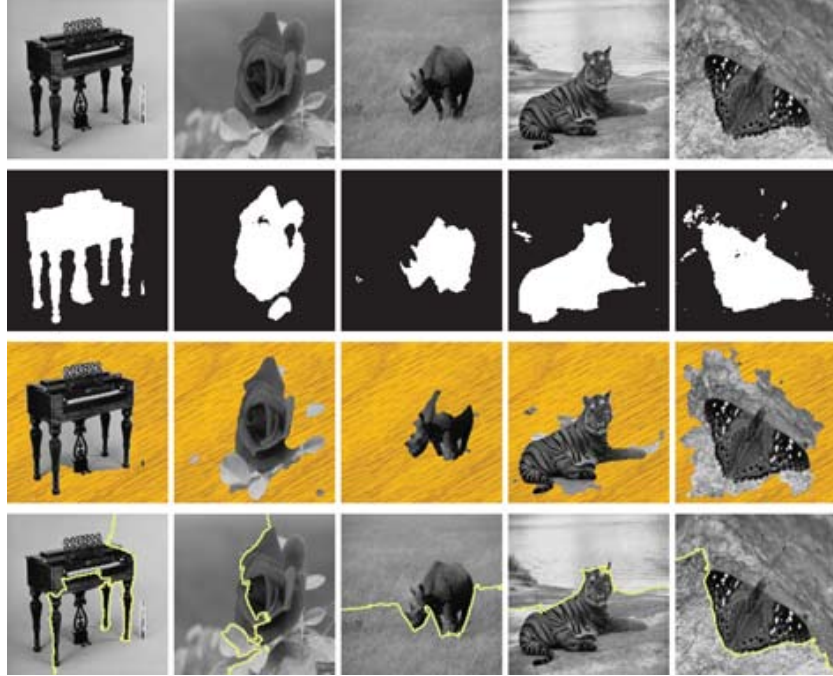


Figure 9: The results of image segmentation. The top row are the segmentation results of active contour [10] on the original images. The bottom row are the results obtained by our Intelligent PDE.

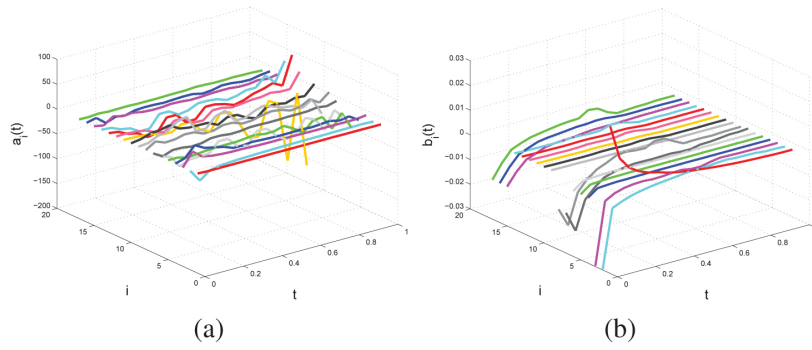


Figure 10: Learnt coefficients of the intelligent PDE system for image segmentation problems. (a) the coefficients  $\{a_j(t)\}_{j=0}^{16}$  for the image evolution. (b) the coefficients  $\{b_j(t)\}_{j=0}^{16}$  for the indicator evolution.

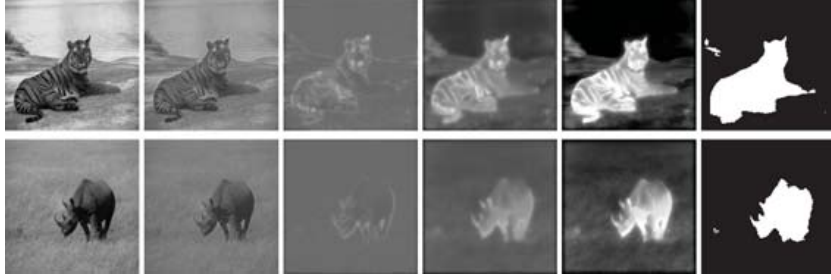


Figure 11: The evolution of the mask maps. For each row, the first image is the input image, the second to the fifth are the mask maps at time  $t = 0.25, 0.50, 0.75, 1.0$ , respectively, and the last image is the final mask map with a threshold of 0.5.

such a problem. As a result, we are unaware of any PDE-based method that can accomplish this task. The existing PDE-based segmentation algorithms always output an “object region” even if the image does not contain the object of interest. In contrast, we will show that as desired our PDEs are able to respond selectively. We choose the “plane” data set in Corel [23]. We select 30 images from this data set as positive samples and also prepare 30 images without the object of interest as negative samples. We also provide their ground truth object masks<sup>6</sup> in order to complete the training data.

In Fig. 12, one can see that our learnt PDEs respond well to the objects of interest (first three images), while the response to images without the objects of interest is relatively low across the whole images (last two images). It seems that our PDEs automatically identify that the concurrent high-contrast edges/junctions/corners are the key features of planes. The above examples show that our learnt PDEs are able to differentiate the object/non-object regions, without requiring the user

---

<sup>6</sup>For the positive samples, we manually segment binary masks as the output images. For the negative samples, the ground truth output masks are all-zero images.



to teach them what features are and what factors to consider. The curves of the learnt coefficients for object detection are shown in Fig. 13.



Figure 12: The results of “plane” detection. The top row are the original images, the first three are images containing planes and the last two are without planes. The bottom row are the detection results of our intelligent PDE for images with and without planes, respectively.

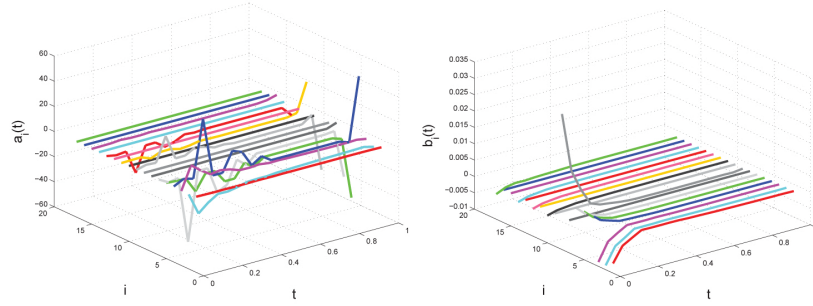


Figure 13: Learnt coefficients of the intelligent PDE system for the plane detection problem. (a) the coefficients  $\{a_j(t)\}_{j=0}^{16}$  for the image evolution. (b) the coefficients  $\{b_j(t)\}_{j=0}^{16}$  for the indicator evolution.

## 5. Conclusion

In this paper, we have presented a framework for using data-based optimal control to learn PDEs as a general regressor to approximate the nonlinear map-

pings of different visual processing tasks. The experimental results on some computer vision and image processing problems show that our framework is promising. However, the current work is still preliminary, so we plan to improve and enrich our work in the following aspects. First, more theoretical issues should be addressed for this PDE system. For example, we will try to apply the Adomian decomposition method [25] to express the exact analytical solution to (1) and then analyze its physical properties. Second, we would like to develop more computationally efficient numerical algorithms to solve our PDE-constrained optimal control problem (4). Third, we will apply our framework to more vision tasks to find out to what extent it works.

#### Appendix A. Proof of Property 2.1

We prove that the coefficients  $\{a_j\}_{j=0}^{16}$  and  $\{b_j\}_{j=0}^{16}$  must be independent of  $(x, y)$ .

*Proof.* We prove for  $F(u, v, \{a_j\}_{j=0}^{16})$  in (1) only. We may rewrite

$$F(u, v, \{a_j\}_{j=0}^{16}) = \tilde{F}(u, v, x, y, t).$$

Then it suffices to prove that  $\tilde{F}$  is independent of  $(x, y)$ .

By the definition of translational invariance, when  $I(x, y)$  changes to  $I(x - x_0, y - y_0)$  by shifting with a displacement  $(x_0, y_0)$ ,  $u(x, y)$  and  $v(x, y)$  will change to  $u(x - x_0, y - y_0)$  and  $v(x - x_0, y - y_0)$ , respectively. So the pair  $u(x - x_0, y - y_0)$  and  $v(x - x_0, y - y_0)$  fulfils (1), i.e.,

$$\begin{aligned} & \frac{\partial u(x - x_0, y - y_0)}{\partial t} \\ & - \tilde{F}(u(x - x_0, y - y_0), v(x - x_0, y - y_0), x, y, t) = 0. \end{aligned}$$

Next, we replace  $(x - x_0, y - y_0)$  in the above equation with  $(x, y)$  and have:

$$\frac{\partial u(x, y)}{\partial t} - \tilde{F}(u(x, y), v(x, y), x + x_0, y + y_0, t) = 0.$$

On the other hand, the pair  $(u(x, y), v(x, y))$  also fulfils (1), i.e.,

$$\frac{\partial u(x, y)}{\partial t} - \tilde{F}(u(x, y), v(x, y), x, y, t) = 0.$$

Therefore,  $\tilde{F}(u, v, x + x_0, y + y_0, t) = \tilde{F}(u, v, x, y, t), \forall (x_0, y_0)$  that confines the input image inside  $\Omega$ . So  $\tilde{F}$  is independent of  $(x, y)$ .  $\square$

## Acknowledgment

This work was partially supported by the grants of the National Science Foundation of China, No. U0935004 and 60873181.

## References

- [1] T. Chen and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. SIAM Publisher, 2005.
- [2] P. Pietro and M. Jitendra, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990.
- [3] S. Osher and L. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM Journal on Numerical Analysis*, vol. 27, pp. 919–940, 1990.
- [4] G. Sapiro, *Geometric partial differential equations and image analysis*. Cambridge University Press, 2001.

- [5] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D.*, vol. 60, pp. 259–268, 1992.
- [6] T. Chan and S. Esedoglu, “Aspects of total variation regularized  $L^1$  function approximation,” *SIAM Journal on Applied Mathematics*, vol. 65, pp. 1817–1837, 2005.
- [7] D. Strong and T. Chan, “Edge-preserving and scale-dependent properties of total variation regularization,” *Inverse Problems*, vol. 19, pp. 165–187, 2003.
- [8] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher, “Fast global minimization of the active contour/snake model,” *Journal of Mathematical Imaging and Vision*, vol. 28.
- [9] A. Chambolle, “Total variation minimization and a class of binary MRF models,” in *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, 2008.
- [10] C. Li, C. Xu, C. Gui, and M. Fox, “Level set evolution without re-initialization: a new variational formulation,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
- [11] X. Gao, B. Wang, D. Tao, and X. Li, “A relay level set method for automatic image segmentation,” *IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics*, vol. 41, pp. 518–525, 2011.
- [12] D. Cheng, *Field and wave electromagnetics (2nd Edition)*. Prentics-Hall, 1989.

- [13] P. Olver, *Applications of Lie groups to differential equations*. Springer-Verlag, 1993.
- [14] D. Kirk, *Optimal control theory: an introduction*. Prentice-Hall, 1971.
- [15] A. Ababnah and B. Natarajan, “Optimal control-based strategy for sensor deployment,” *IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics*, vol. 41, pp. 97–104, 2011.
- [16] J. Lions, *Optimal control systems governed by partial differential equations*. Springer-Verlag, 1971.
- [17] R. Liu, Z. Lin, W. Zhang, and Z. Su, “Learning PDEs for image restoration via optimal control,” in *European Conference on Computer Vision (ECCV)*, 2010.
- [18] A. Jain, “Partial differential equations and finite-difference methods in image processing, part 1,” *Journal of Optimization Theory and Applications*, vol. 23, pp. 65–91, 1977.
- [19] J. Stoer and R. Bulirsch, *Introduction to numerical analysis (2nd Edition)*. Springer-Verlag, 1998.
- [20] N. Mila, “A variational approach to remove outliers and impulse noise,” *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 99–120, 2004.
- [21] J. Parker, *Algorithms for image processing and computer vision*. John Wiley & Sons Inc, 1997.

- [22] D. R. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *International Conference on Computer Vision (ICCV)*, 2001.
- [23] Corel photo library, corel corp., Ottawa, Canada.
- [24] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [25] A. Wazwaz, *Partial differential equations and solitary waves theory*. Springer-Verlag, 2009.